



**The Next Generation of Content
Addressable Memories**

MOSAID Technologies Incorporated
September 1999

The Challenging Market Environment for CAMs

The explosive growth in the use of Intranets and the Internet are taxing LANs to the limit. Overall bandwidth needs are growing quickly. Applications such as L2 bridging and Fast-path L3 routing are now typically implemented in hardware. New requirements, including flow analysis, policy based routing and end-to-end QoS, are rapidly increasing the variety and quantity of lookups required. Every packet must now be examined multiple times and large hardware bridge/router implementations require a forwarding decision to be made every few nanoseconds.

Table sizes are also growing. Even though Classless Inter Domain Routing (CIDR) has managed to slow the growth of internet router table sizes, flow analysis and routing based on flow ID are all creating the need for deeper and wider table sizes in bridges and routers. Currently, table sizes for Internet routing can be as large as 55,000 entries and this is expected to grow at 5,000 entries per year for the foreseeable future. Flow ID table sizes are ultimately limited by the cost of implementation.

Table searching is one of the most time-consuming operations in a router. This bottleneck typically limits the forwarding capability of current routers. Traditionally, Network System Architects have been forced to address these problems using hashing or tree searching algorithms, which are acceptable for moderate performance applications. First generation CAMs did not achieve wide spread adoption for bridging and routing applications because they were slow, expensive, small, and most importantly did not support longest match applications. A new generation of CAMs now addresses these issues.

What is a CAM?

A Content Addressable Memory is a device designed to accelerate any application that requires extremely fast searches of list based data. To best understand what a CAM does, it helps to contrast it to conventional random access memory. Data is stored in memory in specific locations called addresses. When there is a need to retrieve the data, an address is supplied to the memory, which in turn returns the data.

In a CAM, the opposite occurs. Data is supplied to the memory via a special comparand register and the memory returns an address if a corresponding match is found. This enables extremely quick searches. The entire CAM is searched in a single clock cycle and if a match is found, the address returned is used to retrieve data associated with the search string. The associated data is typically stored in a separate, discrete memory in a location specified by the result of the CAM search.

The new generation of CAMs

Now, technology is enabling commercially viable, adequately-sized CAMs. Increasingly, system designers are making the transition from familiar but lower performance solutions to this new breed of CAMs. High performance CAMs, capable of tens of millions of searches per second and cascadable to create table depths of over 500K entries, are now

The Next Generation of Content Addressable Memories

becoming available.

Feature rich, the leading edge CAMs provide functionality including automatic aging, multiple search mask registers, glueless depth expansion, one cycle auto-learning and multiple match priority resolution. These features are all essential for the new generation of switches and routers. Ternary CAMs, able to store 0,1 and "don't care", address the needs of emerging applications such as CIDR, flow analysis, advanced VLAN support and L4 awareness. Per bit ternary CAMs enable longest match searches, in a single cycle, required for high performance routing.

Not limited to routing, CAMs can be used in many application areas including packet classification, compression, cryptography, pattern recognition, and parallel data processing. Larger, faster CAMs offer many advantages as the need for increasing bandwidth continues.

Policy and Flow-based Applications

The various ways in which the Internet is being used, combined with the ever increasing need for bandwidth has created the requirement to assign levels of priorities to different applications. Intending to address this requirement, RSVP and differentiated Services have been developed to allow network providers to give preferential treatment to certain types of traffic. Differentiated Services further implies that traffic types can be distributed into separate classes that can then be given specific priority as they travel through the network.

QoS depends fundamentally on the ability to identify, classify and mark network traffic. Without these basic building blocks, attempts to provide this kind of service will likely not provide the desired behaviour in the network;

QoS is currently being delivered in two basic ways. Policy-based routing is concerned with generic classification of the packet contents for service level determination. Flow-based routing assigns a service level to uniquely identified dialogues (or "Flows"). Policy-based routing can assign QoS based on the "Type of Service" TOS bits as defined by differentiated Services, or add hoc methods (i.e. classifying the application port numbers, FTP Vs. telnet Vs. http).

Flow ID is intended to identify the end-to-end application stream, of which any given packet is a member. Once a packet can be assigned to a flow, it can be forwarded with an associated QoS. A flow can be uniquely identified by a 4-tuple consisting of source IP address, source TCP or UDP port, destination IP address, and destination TCP or UDP port as shown in the table below.

Source IP Address	Source TCP/UDP Port	Destination IP Address	Destination TCP/UDP Port
32 bits	16 bits	32 bits	16 bits

The Next Generation of Content Addressable Memories

This represents a minimum of 96 bits to uniquely identify a flow in IPv4. Adding information such as previous hop further increases the amount or width of data to be processed. IPv6 increases the width considerably with the proposed 128 bit source and destination addresses. Future extension headers such as hop-by-hop options header and routing header will also need to be processed by routers along the transmission path.

Resource ReSerVation Setup Protocol or RSVP is an IP based protocol used for communicating application QoS requirements to intermediate transit nodes in a network. RSVP uses a soft-state mechanism to maintain path and Reservation State in each node in the reservation path. Per flow differentiation allows definition of similar QoS characteristics to particular IP end-to-end sessions.

One major architectural issue in networks that maintain flow-state information centers around the computational and system resource impact on maintaining state information on the potentially thousands of flows established across a diverse network. This is further exacerbated when considering maintaining and manipulating flow information for up to 256,000 active flows in the core of the Internet at any given time. Content addressable memories with the appropriate architecture are ideally suited to address this concern.

CAMs and Network Processors

For some time, centralized software solutions for classification and forwarding lookups have limited router performance. In current architectures, the lack of suitable CAMs has driven the development of multi-probe hash lookup schemes and complex tree-walking architectures to achieve the necessary lookup performance. These solutions did not have enough flexibility to address the current needs for QoS features supported by Flow and Policy based routing. Today, there is a lot of interest in the concept of dedicated processors, called Network Processors (NP), to manage packet classification, admission control and forwarding decisions in a de-centralized fashion.

The Network Processor would use a simplified instruction set and very high speed operation to parse the packet and accelerate table lookup schemes in dedicated memory on a per port, or small aggregate of ports, basis. The NP provides flexibility, but can be somewhat non-deterministic and requires distributing software and databases throughout the system. This can increase system cost.

Ternary CAMs also provide flexibility, allowing the OEM to implement packet classification, QoS determination, and make the forwarding decision in a series of pipelined lookups. This can be accomplished at a lower overall cost, even on a per-port basis. More importantly, the performance of MOSAID's CAMs support much higher throughput allowing more ports to be aggregated together and providing a solution that scales to meet future router needs.

DRAM CAMs

Binary CAMs, capable of storing large tables are now in production and are widely available to System Architects. These CAMs are well suited for a number of applications. New applications such as CIDR, flow analysis, advanced VLAN support and L4 awareness are all creating a need to store "don't care" (X) in addition to ones and zeros. Ternary CAMs, capable of storing 0,1,X in a single cell, are required to provide the benefits of CAMs to these applications.

Binary SRAM CAMs are typically implemented using a 10 transistor design. Early attempts at storing ternary data involved 2 10T SRAM cells and software machinations. Even now, ternary capability in SRAM requires a 16 transistor CAM cell.

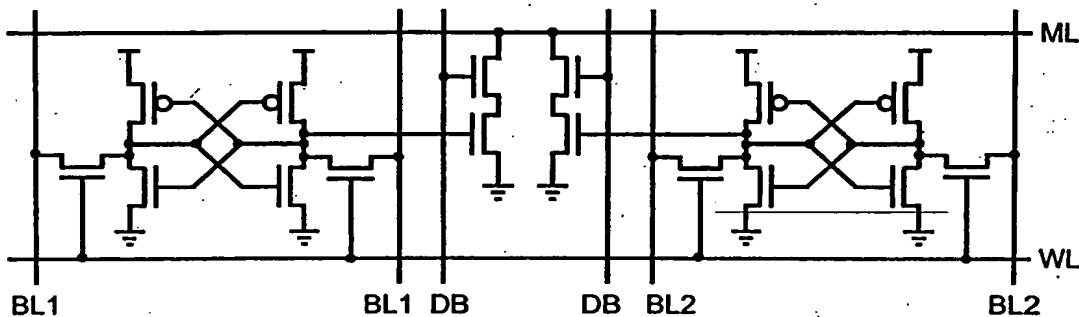


Figure 1: 16T Ternary SRAM Cell

A DRAM CAM enables a denser CAM with a much smaller cell size than competing SRAM based solutions. DRAM CAMs can be designed to be as large or larger than any commercially available monolithic ternary CAM on the market today. DRAM CAM cells can be implemented in just 6 transistors, representing a 2.5 times density advantage of competing SRAM technology.

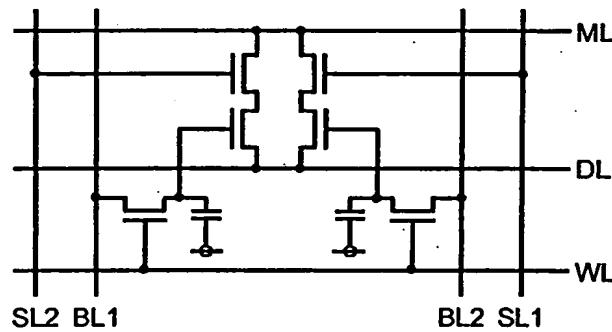


Figure 2: MOSAID 6T Ternary DRAM Cell¹

¹ Patent pending

The Next Generation of Content Addressable Memories

As DRAM CAMs cells are inherently ternary, each 6 transistor cell is capable of storing 0,1,X with no additional overhead. The end result is a lower power, denser and more cost effective approach to CAM implementation. Fixed overhead for the interface, registers and multiple match resolver enable DRAM CAMs to scale to increasingly wide word implementations with only minimal die size penalties. The ability to create cost effective, extremely wide CAMs is critical for emerging applications such as flow analysis, RSVP and IPv6.

A reality of any DRAM-based system is the need for refresh. Designers are typically familiar with DRAM and have experience dealing with this requirement. Most applications tend to require searches in bursts, leaving free cycles in which refresh commands can easily be scheduled. Should system designers wish to avoid refresh scheduling altogether, MOSAID CAMs also provide an auto-refresh mode of operation.

MOSAID DRAM CAMs offer performance that equals or exceeds any currently available SRAM solution. As with many fast, synchronous systems, MOSAID CAMs are pipelined. A pipelined architecture allows optimized read and write throughput. MOSAID has applied this in our CAMs to provide write performance equaling search speeds. This pipelined architecture allows cascading of multiple CAMs for deeper table sizes with no degradation in performance and only one additional cycle latency on the first search.

DRAM CAMs deliver highly scalable, high performance solutions and full ternary capability. DRAM CAMs provide the ultimate flexibility for evolving applications like flow ID and classification, sorting by Layer 2, 3 and 4 address functions.

Highlights of the MOSAID Class-IC CAM

Per Bit Ternary

The inherent ternary capability featured in MOSAID's DRAM-based CAM cell and the multiple match resolver in the Class-IC combine to provide the highest flexibility and support the broadest range of CAM applications.

Of course, router tables requiring longest-match, single cycle, searches can be directly supported in the Class-IC CAM, but the per-bit ternary nature of the Class-IC CAM does not limit implementation to a specific application (e.g. CIDR). The per-bit ternary is quite important for the encoding techniques required to achieve single-search packet classification, but also supports default cases for advanced VLAN support, tailored flood vectors for forwarding, partial matches on Flow IDs, and other innovative applications. The multiple masks and partitioning capability support multiple search scenarios on a single database and multiple databases on the same device.

System-level Features

Associated with each CAM entry are a number of "special bits." These bits are used to encode the type and validity of the entry. In the first generation Class-IC device, bits for Empty, Skip, Permanent and Age have been provided. Empty status is an obvious requirement for updating the table. Skip is important for managing pre-allocated, but empty locations in the CAM and also allows the user to "walk through" a series of

The Next Generation of Content Addressable Memories

multiple matches. Age is a single bit age indication, updated whenever an entry is referred to, provided so that "stale" CAM entries can be purged after a desired interval. Permanent protects an entry against purging due to age. The MOSAID Class-IC CAM can purge all the "old" entries in two clock cycles, minimizing the impact on system performance (throughput).

Learning & Aging support is included in the Class-IC CAM primarily to support Layer-2 Bridging applications in switches. It is also key to the new emerging Flow-based routers, where there may be a need to automatically learn Flows and age them out as well. In Automatic Learning, the user can specify that different masks be used for the search and for the learn operation so that ternary fields can be automatically encoded into the learned data. The Class-IC supports single-cycle learns, sustaining full system throughput.

If Learning & Aging support is not required of the particular application, the "special bits" can be used to partition the CAM into different segments supporting multiple databases without wasting CAM entry bits. Any of the Skip, Permanent and Age bits can be recovered, providing up to eight different partitions.

Interface Considerations.

An important part of a CAM product offering, especially as CAM word widths get wider, is the physical interface offered by the chip. Minimizing the pin-count that is required of an OEM ASIC is very important.

The SDQ (Search Data) interface on Class-IC is bi-directional, so that the ASIC designer can perform both read and write operations with a single port.

A traditional synchronous interface is available for word widths up to 72-bits in MOSAID's first Class-IC CAM. The CAM also features a Dual Data Rate (DDR) interface, commonly available in today's advanced DRAMs. DDR allows the user to clock data in on both edges of the synchronous clock, allowing a 66 MHz search rate to be maintained even when the I/O width is less than the data width. This allows the ASIC designer to work with only 36 SDQ pins for a 72-bit datawidth or support a 144-bit datawidth with only 72 SDQ pins.

Another important concern in minimizing the pin count is to ensure that the OEM can access and manipulate the associated data SRAM without requiring an additional interface on their ASIC. In the Class-IC device this is achieved by providing a specialized instruction that will directly pass an SRAM address from the SDQ pins of the CAM and out onto the MA bus to the SRAM address pins.

Looking Forward – Class-IC Roadmap

DRAM CAM technology offers a higher bit density than is possible using SRAM based solutions. This allows fully ternary CAMs of up to 8Mbits to be quickly developed in existing technology with even larger CAMs becoming available as new fabrication processes become available. Wider CAMs can also easily be implemented in DRAM. As CAMs typically have fixed overhead for logic outside the CAM array, wider DRAM CAMs deliver even higher silicon efficiency.

The Next Generation of Content Addressable Memories

Speed grades can also be easily increased in current technologies. DRAM CAM arrays and associated word-line drivers can be architected to provide significantly higher performance with only a minor reduction in density and increased power consumption. DRAM CAMs can approach the random read/write performance of SRAM solutions and achieve higher search throughputs (133MHz) while continuing to retain a considerable size advantage over SRAM.

Innovative techniques such as the SDRAM 533 MHz bus interface could be used to dramatically reduce the interface pin count to the OEM ASIC. Other advanced techniques could be used to further improve propagation delays and setup times at the interface.

Specialized features can also be easily developed using DRAM CAM technology. Although a DRAM CAM cell is inherently ternary, simplifying the interface for communicating the ternary mask through enumerating the number of significant bits could be particularly useful for specific applications like CIDR. While such a device would be less flexible, it would be optimized to the application.

The pipelined synchronous architecture of the current MOSAID Class-IC device has enabled very high throughput for the current and future generations of OEM systems. The pipelined mode and high-speed synchronous interfaces with DLL clock schemes will allow the MOSAID Class-IC product line to keep pace with the multi-Gigabit and Terabit router/switch architectures currently on the drawing board. The pipelined architecture ensures no performance degradation with multiple Class-IC devices cascaded together. Pipelining can also allow seamless interface to SSRAM for associated data, and "early" status flags to enable conditional processing, all without any penalty to sustained throughput.

Looking forward, the system-level features of the MOSAID Class-IC family will continue to evolve to meet requirements created by emerging applications. Future releases will include features and functionality in such areas as simultaneous multiple word width support, "validity" bits on output, highly flexible partitioning, greater age granularity and special features for "flow" aging (including limited purges), automatic sorting of entries for longest match support, and internal associated data RAM.

For over 25 years, MOSAID Semiconductor has been at the forefront of DRAM design. MOSAID Semiconductor has been instrumental in the development of the last 9 generations of commodity DRAM and has developed high performance, specialized DRAM interfaces including DDR and SLIO. MOSAID is committed to applying this experience to provide industry leading CAM technology to alleviate bottlenecks and accelerate networking applications.